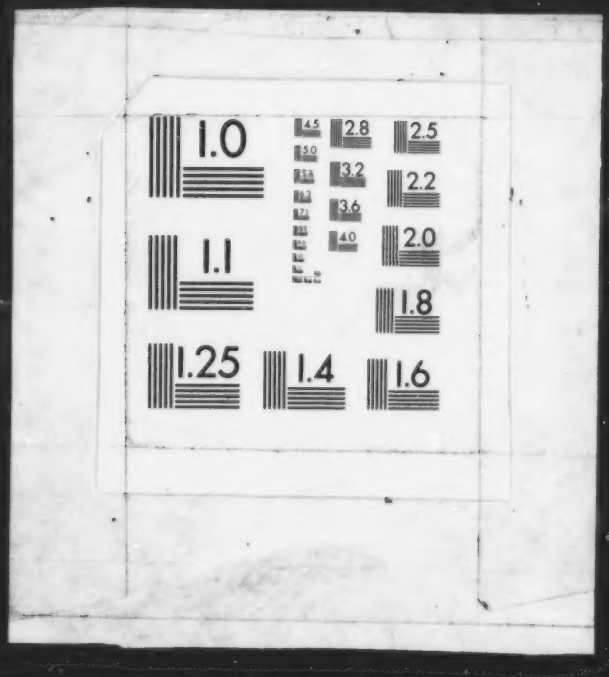
1 OF 1
PB 86
121209



PB86-121209

Estimates of Analytical Accuracy, Repeatability, and Reproducibility from EPA (Environmental Protection Agency) National Coal Audit Program Data

(U.S.) Environmental Monitoring Systems Lab.
Research Triangle Park, NC

Oct 85

U.S. Department of Commerce
National Technical Information Service
NTIS

PB86-121209

EPA/600/D-85/229
October 1985

Estimates of Analytical Accuracy, Repeatability, and Reproducibility from EPA National Coal Audit Program Data

by

Thomas J. Logan,[1] Samuel S. Baker, Jr.,[2] Ellen W. Streib,[1]
M. Rodney Midget,[1] and W. S. Pitts, III[2]

[1]U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

[2]Kilkelly Environmental Associates
Raleigh, NC

EPA Contract No.
68-02-4125

Project Officer
Thomas J. Logan

Environmental Monitoring Systems Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA/600/D-85/229		PB86 121209/AS

4. TITLE AND SUBTITLE
Estimates of Analytical Accuracy, Repeatability, and Reproducibility from EPA National Coal Audit Program Data

5. REPORT DATE
October 1985

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
T. J. Logan,* S.S. Baker, Jr.,**, E.W. Streib,* M.R. Midget,* and W.S. Pitts,III**

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
*USEPA, Research Triangle Park, NC 27711
**Kilkelly Environmental Associates, Raleigh, NC 27622

10. PROGRAM ELEMENT NO.

11. CONTRACT/GRANT NO.
QA-V 68-02-4125

12. SPONSORING AGENCY NAME AND ADDRESS
U.S. Environmental Protection Agency
Environmental Monitoring Systems Laboratory
Quality Assurance Division/Source Branch (MD-77A)
Research Triangle Park, North Carolina 27711

13. TYPE OF REPORT AND PERIOD COVERED

14. SPONSORING AGENCY CODE
EPA/600/08

15. SUPPLEMENTARY NOTES

16. ABSTRACT
The accuracy, repeatability, and reproducibility of coal sulfur, moisture, and Btu determinations were estimated from results reported during four rounds of EPA's National Coal Audit Program (NCAP). For seven out of eight groups of audit samples (analyzed by 63 to 120 laboratories), the averages of reported sulfur results were up to 5.08 percent low relative to audit sample sulfur content target values. From 88.9 percent to 96.3 percent of duplicated sulfur determinations met ASTM repeatability criteria. Mean differences between reported Btu determinations and Btu target values ranged from -0.71 percent to +1.27 percent of the target values. From 83.6 percent to 91.9 percent of duplicated Btu determinations were within 50 Btu. Over the eight groups of audit samples, from 56.7 percent to 84.2 percent of laboratory-to-laboratory differences between averaged, within-laboratory duplicate sulfur determinations were within 0.05 percent sulfur for audit samples with sulfur contents less than 2 percent, or within 0.10 percent sulfur for audit samples with sulfur contents greater than 2 percent. From 44.0 percent to 70.2 percent of laboratory-to-laboratory differences in Btu results were within 100 Btu. Overall, the data examination indicates that in instances where coal analysis data are used for determining source-specific SO_2 emissions, there may be a need for enhanced analytical quality assurance and quality control measures.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	20 16
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ESTIMATES OF ANALYTICAL ACCURACY, REPEATABILITY, AND REPRODUCIBILITY FROM EPA NATIONAL COAL AUDIT PROGRAM DATA

This paper summarizes coal analysis data reported to the U.S. Environmental Protection Agency under the agency's voluntary National Coal Audit Program. The accuracy, repeatability and reproducibility of participating laboratories' results are summarized for sulfur, moisture and Btu content of audit samples distributed during four recent rounds of the analytical audit program.

Introduction - Overview of EPA's National Coal Audit Program

In July 1981, the U.S. Environmental Protection Agency (EPA) Environmental Monitoring Systems Laboratory/Quality Assurance Division (EMSL/QAD) initiated a voluntary, semi-annual National Coal Audit Program (NCAP). The NCAP affords participating laboratories the opportunity to analyze coal samples of known constituent contents (audit samples). EMSL/QAD informs participants of how their analytical results compare to results reported by other laboratories.

In each NCAP "round", participating laboratories are supplied two audit samples. Participants are asked to analyze both samples, in duplicate, and report measured content of sulfur, moisture, Btu and ash. Identification of analytical methods is requested. In many ways, the NCAP is similar to a multi-round coal audit "round robin," in which the sulfur, moisture, Btu and ash content of samples varies from round to round. Throughout this paper, the term "round" is used to designate the semi-annual date on which a set of two audit samples was distributed to NCAP participants.

Audit samples are prepared by commercial coal testing laboratories. A "target value" for each of the four coal constituents of interest is assigned by EMSL/QAD to each lot of audit samples. Target values are based on supplier analyses of the lot of audit samples, and are verified by independent second- and third-party analyses of random samples from each lot. Second- and third-party verification is performed by commerical laboratories under contract to EMSL/QAD.

For information about participating in NCAP, contact Ms. Ellen Streib, Environmental Monitoring Systems Laboratory [MD 77A], U. S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711.

The following sections of this paper (1) define various terms relating to the accuracy and precision of coal analyses, (2) describe data analysis methodologies applied in examining NCAP data, and (3) summarize the authors' findings and conclusion regarding the accuracy and precision of coal sulfur, moisture and Btu data reported by NCAP participants during four recent rounds of the program.

Definitions and Data Analysis Methodologies

Various terms used in this paper are defined as follows:

Accuracy -- signed difference between average of reported duplicate results and target value

Repeatability -- absolute value of within-laboratory difference between duplicate determinations

Reproducibility -- absolute value of laboratory-to-laboratory differences among means of within-laboratory duplicate determinations.

Figure 1 illustrates the approach taken for determining accuracy, repeatability and reproducibility of analytical results in the NCAP data base. Sulfur data are used for the example illustrated in Figure 1.

Columns designated by an asterisk (*) in Figure 1 show examples of the types of information in the data base. The remaining columns show values computed from the basic data. For example, duplicate determinations (columns 4 and 5) were averaged (column 7) for each laboratory. As shown in column 8, differences were computed between the values in column 7 and the NCAP audit sample target value (column 3). The column 8 values were used in examining accuracy.

As shown in column 6, the absolute value of the differences between within-laboratory duplicate determinations were computed. These differences were used in examining repeatability. Repeatability values for all labs were arranged in cumulative frequency distributions from which it is possible to determine the number of individual labs, or percentage of all labs, that achieved a given precision of analysis.

Referring once more to Figure 1, columns 9 through 11 illustrate the approach taken in computing reproducibility values. The columns show that all combinations (used once) of laboratory-to-laboratory differences between the average of within-laboratory duplicate determinations were computed. The absolute value of the between-laboratory differences (i.e., reproducibility values) were also arranged in cumulative frequency distributions from which it is possible to assess between-laboratory precision.

Repeatability and reproducibility criteria from American Society for Testing and Materials (ASTM) Standard Methods for determining sulfur, Btu, and moisture content of coal (i.e., ASTM Methods D 3177-82, D 2015-77 and D 3173-73, respectively) were used as yard sticks for assessing the precision of analytical results in the NCAP data base. The percentage of results meeting those criteria (determined from the repeatability and reproducibility frequency distributions) are reported in this article.

Data Base Size and Screening

Because EPA standards for SO_2 emissions from coal-burning facilities are expressed in terms of lb $SO_2/10^6$ Btu, EMSL/QAD is interested in the sulfur to Btu ratio computed from coal sulfur and coal Btu measurements. Since the focus of this paper is on the accuracy and precision of sulfur and Btu measurements, ash data in the NCAP data base are not discussed. In order to cover an appropriately wide range of sulfur and Btu content samples, data were examined for the March and September 1982, and March and September 1983 NCAP rounds.

During those four rounds, participants reported a total of 4554 sulfur, moisture or Btu values. The vast majority (4422) were results

of duplicate determinations; however, in a few (132) instances, some laboratories reported only a single analytical value for some constituents in some samples. Because single analytical results are unsuitable for examination by the methodology described previously, the 132 unpaired values were not considered in this study.

The 4422 values (2211 duplicate determinations) were screened to identify paired observations exhibiting extreme discrepancy from the corresponding target value. Differences between the average of duplicate determinations and the target value were plotted as distributions, roughly centered around zero, and values on the tails of the distributions were examined. If the average of duplicate analyses differed from the target by more than twice the amount that the next closest average differed from the target, then the average was classified as an outlier. Less than 0.8 percent of the paired observations were classified as outliers. The following sections of this paper are differentiated according to coal constituent. Accuracy, repeatability and reproducibility of analytical results in the NCAP data base are discussed by constituent.

Results

Sulfur Accuracy. The mean of reported results of sulfur analyses was less than the target value for seven out of eight sets of audit samples analyzed by participants in the four NCAP rounds. Over the eight sets of audit samples, negative mean differences from the target value ranged from -0.68 percent to -5.08 percent of the target. In absolute terms, the negative mean differences ranged from -.005 percent to -.099 percent sulfur. For one set of audit samples (in the 9-82 round), participants' results exceeded the target value by an average of +0.018 percent sulfur, or +1.48 percent of the target value.

Sulfur Repeatability. Over the eight sets of audit samples, from 88.9 percent to 96.3 percent of reported results met ASTM repeatability criteria for determination of coal sulfur. That is, from 88.9 percent to 96.3 percent of the within-laboratory differences between duplicate sulfur determinations were less than or equal to 0.05 or 0.10 percent sulfur, where the 0.05 percent criterion applies to samples with a sulfur content of less than 2 percent, and the 0.10 percent criterion applies to samples with a sulfur content of 2 percent or greater.

Sulfur Reproducibility. Taking into consideration all laboratory-to-laboratory comparisons between averaged, within-laboratory duplicate determinations over the eight sets of audit samples, from 56.7 percent to 84.2 percent of reported results met the ASTM reproducibility criteria for coal sulfur determination. The accuracy, repeatability and reproducibility of sulfur results in the NCAP data base are summarized in Table I.

Btu Accuracy. Over the eight sets of audit samples, mean differences from the target value ranged from -90 Btu to +143 Btu. Expressed as a percentage of the target value, mean differences ranged from -0.71 percent to +1.27 percent of the target value. For one set of audit samples, the mean difference from the target value was less than 0.01 percent of the target. These results are summarized in the "Accuracy" section of Table II.

Btu Repeatability. Over the four NCAP rounds, from 83.6 percent to 91.9 percent of results met the ASTM repeatability criterion of 50 Btu maximum expected difference between within-laboratory duplicate determinations.

Btu Reproducibility. From 44.0 percent to 70.2 percent of between-laboratory comparisons met the ASTM reproducibility criterion of 100 Btu maximum expected difference between laboratories, based on averages of within-laboratory duplicate determinations. Percentages of results achieving repeatability and reproducibility criteria are presented in Table II.

Moisture Accuracy. Over the eight sets of NCAP samples, the mean difference between reported results and the target value ranged from -2.63 percent to +6.02 percent of the target value. In absolute terms, the mean differences ranged from -0.161 to +0.159 percent moisture. A moisture data accuracy summary is presented in Table III.

Moisture Repeatability. From 94.2 percent to 100 percent of results met ASTM repeatability criteria. That is, over the eight sets of samples, from 94.2 percent to 100 percent of within-laboratory differences between duplicate determinations were less than or equal to 0.20 percent moisture or 0.30 percent moisture, where the 0.20 percent criterion applies to samples with a moisture content of less than 5 percent and the 0.30 percent criterion applies to samples with a moisture content in excess of 5 percent.

Moisture Reproducibility. Over the eight sets of samples, from 43.9 percent to 65.8 percent of laboratory-to-laboratory differences, based on within-laboratory duplicate determinations, were within ASTM reproducibility criteria. Moisture data repeatability and reproducibility summaries are also presented in Table III.

Round-to-Round Consistency of Laboratory Performance

Experience in other EPA audit programs has shown that some audit participants (or groups of participants) repeatedly out-perform the mean of all participants; and conversely, some participants (or groups of participants) repeatedly under-perform relative to the group as a whole. This observation was examined in regard to the accuracy and repeatability of sulfur and Btu data reported in NCAP.

In assessing the round-to-round consistency of performance by NCAP participants, NCAP rounds were identified for which the sulfur and Btu contents of samples were similar. Similarity in constituent content was considered important in order to eliminate the potentially confounding influence of different constituent levels on the round-to-round consistency of performance by a group of laboratories. The following tabulation lists rounds with similar sulfur and Btu target values.

Target Values for Constituents in Selected Rounds

Round	Sulfur	Btu
3-83	0.74 2.98	14165
9-83	0.74 2.95	14059

Subsets of NCAP data were identified consisting of data reported by laboratories that had participated in both rounds during which constituent content was similar. In examining round-to-round consistency, the accuracy and repeatability of results reported by all participants common to both rounds was determined separately for each of the two rounds. Groups of laboratories achieving the "most accurate," "least accurate," "most repeatable" and "least repeatable" results were identified. Performance of the groups of laboratories was then compared from round-to-round. Findings were as follows.

Consistency of Sulfur Accuracy. Round-to-round consistency in the accuracy of sulfur results was examined for a set of low sulfur and a set of high sulfur audit samples.

Low sulfur. There was a high degree of round-to-round consistency, particularly for the group of least accurate laboratories. Indicative of this consistency, the five least accurate laboratories in the 3-83 round missed the target value by an average of 0.154 percent sulfur (20.8 percent of the target). In the 9-83 round, those same five laboratories missed the target by 0.133 percent sulfur (18.0 percent of the target).

High sulfur. As with accuracy for low sulfur samples, there was a high degree of round-to-round consistency in the accuracy of high sulfur samples analyses -- particularly for those laboratories that missed the target values by the greatest amounts.

Consistency of Sulfur Repeatability. Somewhat surprisingly, for both low sulfur and high sulfur content audit samples, examination indicated a lack of round-to-round consistency in repeatability of sulfur analyses. Groups of laboratories that achieved either the best or the worst repeatability in a given round achieved essentially average repeatability in the other round in which they analyzed similar sulfur content samples.

Consistency of Btu Accuracy. Accuracy of Btu analysis was not found to be consistent from round to round. For example, the five most accurate laboratories in the 3-83 round were within 2.7 Btu of the target value. However, those same five labs were only within 73 Btu of the target in the 9-83 round. The inconsistency also seems to hold for groups of laboratories that reported relatively inaccurate results in a given round. For example, the five least accurate laboratories in the 3-83 round missed the target by an average of 417.5 Btu/lb. However, in the 9-83 round, those same five labs missed the target by 94.6 Btu/lb -- only 21.6 Btu/lb more than the most accurate labs in the 3-83 round. Thus, accuracy in Btu determination was not found to be particularly consistent from round to round.

Consistency of Btu Repeatability. Examination indicated that laboratories achieving superior Btu repeatability in a given round tended to achieve superior repeatability in another round when the Btu contents of the samples were similar, and that laboratories achieving inferior repeatability in a given round tended to achieve inferior repeatability in another round when the Btu contents of the sample were similar. Thus, the repeatability of Btu results by small groups of "superior" or "inferior" laboratories was found to be consistent from round to round.

Round-to-round consistency of sulfur, moisture and Btu accuracy and repeatability is summarized in Table IV.

Sulfur to Btu Ratio

Because sulfur dioxide emission standards are specified in terms of lb $SO_2/10^6$ Btu, and because coal analysis data can, in some cases, be used to determine emission levels, EMSL/QAD is particularly interested in the accuracy (and precision) of sulfur to Btu ratios. For seven out of eight sets of audit samples, the mean lb $SO_2/10^6$ Btu ratio computed from participants' analytical results was less than the NCAP target value. Negative mean differences ranged from -0.48 percent to -6.12 percent of the target value.

Two sets of audit samples were "compliance coals" i.e., coals with less than the equivalent of 1.2 lb $SO_2/10^6$ Btu, when burned. NCAP participants' results exhibited smaller mean percentage differences from the target value for these samples than for the six higher lb $SO_2/10^6$ Btu samples. The accuracy of NCAP data in terms of lb $SO_2/10^6$ Btu is summarized in Table V.

Another way of looking at the accuracy of lb $SO_2/10^6$ Btu is on the basis of the percentage of results that missed the target value by more than $\pm$ 10 percent. In particular, EMSL/QAD is interested in the percentage of results that underestimate coal $SO_2/10^6$ Btu ratio by more than 10 percent. Table VI summarizes NCAP data in this context. It can be seen from Table VI that over the eight sets of audit samples, from 5.6 percent to 16.2 percent of results underestimated lb $SO_2/10^6$ Btu ratio by more than 10 percent.

Summary and Conclusions

Considering that analytical repeatability (i.e., the within-laboratory difference between duplicate determinations) is controllable within each individual laboratory, the authors were surprised to observe that in many (13 out of 16) instances less than 95 percent of reported sulfur and Btu values met applicable ASTM repeatability criteria for individual sets of audit samples. These results are in stark contrast to the precision of moisture determination, which was found to meet ASTM criteria for six out of eight sets of audit samples. The contrast in results can be seen in Table IV.

Based on the similarity in accuracy of sulfur and lb $SO_2/10^6$ Btu ratio values in the NCAP data base (see Table IV) analytical accuracy in sulfur determination was found to be the most important factor contributing to the accuracy of lb $SO_2/10^6$ Btu ratio values calculated from coal analysis data. Although NCAP participants, as a group, estimated the lb $SO_2/10^6$ Btu ratio of audit samples very accurately -- particularly for low sulfur compliance coal samples -- nonetheless, a considerable percentage of participating laboratories underestimated the lb $SO_2/10^6$ Btu ratio by more than 10 per-cent. This observation may point to a need for enhanced analytical quality assurance and quality control in instances where coal analysis data are used for determining source-specific SO_2 emissions.

Table 1. Sulfur accuracy, repeatability, and reproducibility summary

Sulfur Target Value (TV) Round Date		0.74 (3-83)	0.74 (9-83)	1.22 (9-82)	1.30 (3-82)	2.85 (3-82)	2.95 (9-83)	2.98 (3-83)	3.22 (9-82)
Accuracy									
Mean of Participants Results		0.734	0.735	1.238	1.234	2.806	2.851	2.881	3.177
Difference (Δ) from TV		-0.006	-0.005	+0.018	-0.066	-0.044	-0.099	-0.099	-0.043
Δ as a % of TV		-0.81%	-0.68%	+1.48%	-5.08%	-1.54%	-3.36%	-3.36%	-1.34%
Repeatability[f]									
Repeatability Criteria[a]	0.05[b]	<u>89.8</u>	<u>95.3</u>	<u>88.9</u>	<u>90.2</u>	77.0	80.7	76.9	75.7
	0.10[c]	99.2	99.5	95.8	97.0	<u>90.2</u>	<u>96.3</u>	<u>94.0</u>	<u>95.9</u>
Reproducibility[f]									
Reproducibility Criteria[a]	0.10[d]	<u>84.2</u>	<u>72.5</u>	<u>56.7</u>	<u>68.6</u>	37.7	41.3	49.7	39.6
	0.20[e]	98.9	94.6	82.2	90.2	<u>62.0</u>	<u>67.7</u>	<u>72.0</u>	<u>65.8</u>

a From ASTM Standard Method D 3177-82.

b Repeatability criterion for Eschka and Bomb Washing analyses of coals containing <2% sulfur, and High Temperature Combustion Acid/Base Titration analyses of coals irrespective of sulfur content.

c Repeatability criterion for Eschka and Bomb Washing analyses of coals containing ≥2% sulfur.

d Reproducibility criterion for Eschka and Bomb washing analyses of coals containing ≤2% sulfur.

e Reproducibility criterion for Eschka and Bomb washing analyses of coals containing ≥ 2% sulfur.

f Percentage of results achieving criteria.

NOTE: In this and subsequent tables, values relevant to ASTM criteria are underlined. Other values are included for overall comparison.

Table 11. Btu accuracy, repeatability, and reproducibility summary

Btu Target Value (TV) Round Date		9322 (3-82)	11278 (3-82)	12277 (9-82)	12470 (9-83)	12719 (3-83)	12932 (9-82)	14059 (9-83)	14165 (3-83)
Accuracy									
Mean of Participants Results		9411	11421	12244	12569	12629	12893	14059	14106
Difference (△) from TV		+89	+143	-33	+99	-90	-39	<1	-59
△ as a % of TV		+0.95%	+1.27%	-0.27%	+0.79%	-0.71%	-0.30%	<0.01%	-0.42%
Repeatability[b]									
Repeatability Criterion[a]:	50 Btu	91.9	91.8	83.6	90.8	84.6	85.7	90.7	89.7
Reproducibility[b]									
Reproducibility Criterion[a]:	100 Btu	70.2	61.7	47.6	45.9	59.5	44.0	63.6	66.8

[a] From ASTM Standard Method D 2015-77.

[b] Percentage of results achieving criterion.

Table III. Moisture accuracy, repeatability, and reproducibility summary

Moisture Target Value (TV) Round Date		0.96 (3-83)	1.62 (9-83)	1.73 (3-82)	2.05 (3-83)	2.11 (9-82)	2.48 (9-82)	5.75 (9-83)	6.13 (3-83)
Accuracy									
Mean of Participants Results		1.015	1.586	1.718	2.006	2.237	2.559	5.909	5.969
Difference (△) from TV		+0.055	-0.034	-0.012	-0.044	+0.127	+0.079	+0.159	-0.161
△as a % of TV		+5.73%	-2.10%	-0.69%	-2.15%	+6.02%	+3.19%	+2.77%	-2.63%
Repeatability[f]									
Repeatability Criteria[a]	0.20%[b]	<u>100.0</u>	<u>100.0</u>	<u>100.0</u>	<u>94.2</u>	<u>96.0</u>	<u>96.2</u>	93.6	90.8
	0.30%[c]	100.0	100.0	100.0	99.7	97.7	97.9	<u>98.2</u>	<u>94.2</u>
Reproducibility[f]									
Reproducibility Criteria[a]	0.30[d]	<u>59.0</u>	<u>59.3</u>	<u>52.0</u>	<u>64.8</u>	<u>47.5</u>	<u>43.9</u>	42.2	46.9
	0.50[e]	79.3	77.3	71.9	81.5	66.7	64.5	<u>60.3</u>	<u>65.8</u>

a From ASTM Standard Method D 3173-73.

b Repeatability criterion for coal of moisture content <5%.

c Repeatability criterion for coal of moisture content >5%.

d Reproducibility criterion for coal moisture content <5%.

e Reproducibility criterion for coal moisture content >5%.

f Percentage of results achieving criteria.

NOTE: Values relevant to ASTM criteria are underlined. Other values are included for overall comparison.

Table IV. Overall summary of accuracy, repeatability, reproducibility and consistency for sulfur, moisture, Btu and sulfur to Btu ratio

	BTU	SULFUR		MOISTURE		Lb $SO_2/10^6$ Btu	
		<2%	>2%	<5%	>5%	<1.20	>1.20
Accuracy							
Mean % Difference from Target Value[a]	-0.71% to +0.95%	-5.08% to +1.48%	-3.36% to -1.34%	-2.15% to +6.02%	-2.63% to +2.77%	-0.66% to -0.48%	-6.12% to +2.01%
Narrowest Range for 80% of Data[b]	-0.9% to +0.2%	-11.5% to +1.8%	-10.4% to +0.8%	-22.9% to +10.2%	-12.1% to +4.4%	-7.2% to +6.2%	-9.1% to +1.6%
Widest Range for 80% of Data[b]	-2.2% to +0.7%	-14.4% to +7.6%	-11.6% to +5.0%	-43.8% to +32.3%	-9.2% to +10.6%	-15.6% to +7.2%	-9.9% to +11.7%
Repeatability[a]							
% of Data Meeting ASTM Criterion	83.6% to 91.9%	88.9% to 95.3%	90.2% to 96.3%	96.0% to 100%	94.2% to 98.2%	3.0% to 4.1%[c]	1.8% to 2.3%[c]
Reproducibility[a]							
% of Data Meeting ASTM Criterion	44.0% to 70.2%	56.7% to 84.2%	62.0% to 72.0%	43.9% to 64.8%	60.3% to 65.8%	12.0% to 16.7%[d]	8.3% to 14.9%[d]
Round to Round Consistency							
Accuracy	I[f]	C[e]	C[e]	I[f]	--	--	--
Repeatability	C[e]	I[f]	I[f]	I[f]	--	--	--

[a] Range for all samples.

[b] Range (80%) of reported data for a single set of samples.

[e] Consistent

[f] Inconsistent

[c] There is no applicable ASTM criterion. Repeatability is expressed as a percent of TV, at 80th percentile, for best and worst set out of the eight sets of samples.

[d] There is no applicable ASTM criterion. Reproducibility is expressed as a percent of TV, at 80th percentile, for best and worst set out of the eight sets of samples.

Table V. Lb $SO_2/10^6$ Btu data accuracy summary

Lb $SO_2/10^6$ Btu TV[a] Round Date	1.045 (3-83)	1.054 (9-83)	1.987 (9-82)	2.305 (3-82)
Mean of Participants Results	1.040	1.047	2.027	2.164
Difference (△) from TV	-0.005	-0.007	+0.040	-0.141
△ as a % of TV	-0.48%	-0.66%	+2.01%	-6.12%
Range for 80% of Data	0.97 to 1.11	0.89 to 1.13	1.79 to 2.22	2.03 to 2.33
Range of △'s for 80% of Data	-.075 to +.065	-.164 to +.076	-.197 to +.233	-.275 to +.025
Range of △'s as a % of TV	-7.18% to +6.22%	-15.56% to +7.21%	-9.91% to +11.73%	-11.93% to +1.08%

Lb $SO_2/10^6$ Btu TV[a] Round Date	4.686 (3-83)	4.731 (9-83)	4.980 (9-82)	6.115 (3-82)
Mean of Participants Results	4.559	4.539	4.926	5.965
Difference (△) from TV	-0.127	-0.192	-0.054	-0.150
△ as a % of TV	-2.71%	-4.06%	-1.08%	-2.45%
Range for 80% of Data	4.26 to 4.76	4.27 to 4.82	4.60 to 5.18	5.45 to 6.36
Range of △'s for 80% of Data	-.426 to +.074	-.461 to +.089	-.380 to +.200	-.665 to +.245
Range of △'s as a % of TV	-9.09% to +1.58%	-9.74% to +1.88%	-7.63% to +4.02%	-10.87% to +4.01%

[a] TV = Target Value

Table VI. Lb $SO_2/10^6$ Btu accuracy - percent of data more than ± 10 percent different from target value

Lb $SO_2/10^6$ Btu TV[a] Round Date	1.045 (3-83)	1.054 (9-83)	1.987 (9-82)	2.305 (3-82)
Mean of Participants Results	1.040	1.047	2.027	2.164
Difference (△) from TV	-0.005	-0.007	+0.040	-0.141
△as a % of TV	-0.48%	-0.66%	+2.01%	-6.12%
% Data <0.9 x TV	5.6%	15.7%	9.8%	16.2%
% Data >1.1 x TV	6.7%	8.2%	12.0%	0.0%
% Data Outside ± 10% TV	12.3%	23.9%	21.8%	16.2%

Lb $SO_2/10^6$ Btu TV[a] Round Date	4.686 (3-83)	4.731 (9-83)	4.980 (9-82)	6.115 (3-82)
Mean of Participants Results	4.559	4.539	4.926	5.965
Difference (△) from TV	-0.127	-0.192	-0.054	-0.150
△as a % of TV	-2.71%	-4.06%	-1.08%	-2.45%
% Data <0.9 x TV	7.5%	9.1%	7.5%	12.2%
% Data >1.1 x TV	1.2%	1.0%	1.2%	3.2%
% Data Outside ± 10% TV	8.7%	10.1%	8.7%	15.4%

[a] TV = Target Value

Figure 1. Illustration of Approach for Determining Accuracy, Repeatability and Reproducibility From NCAP Data

		Round 1 Sample 1 (e.g., 1.0% S)											Round 1 Sample 2 (e.g., 2% S)
(1)* Laboratory	(2)* Method	(3) T	(4)* Anal 1 (A_1)	(5)* Anal 2 (A_2)	(6) $\lvert A_1 - A_2 \rvert$ (Δ)	(7) $(A_1 + A_2)/2$ ($\overline{A}_1$)	(8) $\overline{A}_1 - T$	(9) $\lvert \overline{A}_A - \overline{A}_1 \rvert$	(10) $\lvert \overline{A}_B - \overline{A}_1 \rvert$	(11) $\lvert \overline{A}_C - \overline{A}_1 \rvert$	. . .	(n) $\lvert \overline{A}_G - \overline{A}_1 \rvert$	
Lab "A"	X	1.0	1.02	1.04	0.02	1.03	0.03	--	--	--	. . .	--	
Lab "B"	X	1.0	0.98	1.00	0.02	0.99	-0.01	0.04	--	--	. . .	--	
Lab "C"	X	1.0	0.98	1.00	0.02	0.99	-0.01	0.04	0	--	. . .	--	
Lab "D"	Y	1.0	1.00	1.00	0	1.00	0	0.03	0.01	0.01	. . .	--	
Lab "E"	Y	1.0	1.02	1.06	0.04	1.04	0.04	0.01	0.05	0.05	. . .	--	
Lab "F"	Y	1.0	0.96	0.98	0.02	0.97	-0.03	0.06	0.02	0.02	. . .	--	
Lab "G"	Z	1.0	0.98	1.02	0.04	1.00	0	0.03	0.01	0.01	. . .	--	
Lab "H"	Z	1.0	1.00	1.04	0.04	1.02	0.02	0.01	0.03	0.03	. . .	0.02	
Lab "A"	X				Repeatability		Accuracy	Reproducibility					
Lab "B"	X												
Lab "C"	X												
Lab "E"	Y												
Lab "G"	Z												
Lab "H"	Z												
Lab "J"	X												
Lab "K"	Y												

OTHER ROUNDS ↓

Notes:

*Designates information coded in data base. Other column values were computed.

X, Y, Z designate alphanumeric method codes.

T = NCAP target value.

$\overline{A}_A = \overline{A}_1$ for Lab A; $\overline{A}_B = \overline{A}_1$ for Lab B, etc.